# wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 02:49 PM BST

| | | |
|---|---|---|
| PDB ID | : | 1HY8 |
| Title | : | SOLUTION STRUCTURE OF B. SUBTILIS ACYL CARRIER PROTEIN |
| Authors | : | Xu, G.-Y.; Tam, A.; Lin, L.; Hixon, J.; Fritz, C.C.; Power, R. |
| Deposited on | : | 2001-01-18 |

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

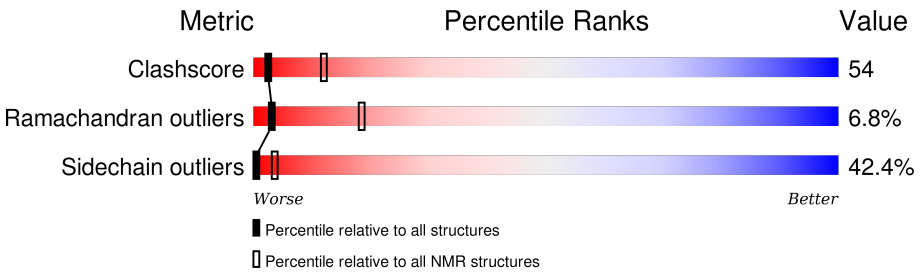
| | | |
|---|---|---|
| Cyrange | : | Kirchner and Güntert (2011) |
| NmrClust | : | Kelley et al. (1996) |
| MolProbity | : | 4.02b-467 |
| Mogul | : | unknown |
| Percentile statistics | : | 20151230.v01 (using entries in the PDB archive December 30th 2015) |
| RCI | : | v_1n_11_5_13_A (Berjanski et al., 2005) |
| PANAV | : | Wang et al. (2010) |
| ShiftChecker | : | rb-20027457 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | rb-20027457 |

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 84%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




| Metric | Whole archive (#Entries) | NMR archive (#Entries) |
|---|---|---|
| Clashscore | 114402 | 11133 |
| Ramachandran outliers | 111179 | 9975 |
| Sidechain outliers | 111093 | 9958 |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 76 | 17% 66% 16% • |

## 2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

# 3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1158 atoms, of which 566 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called ACYL CARRIER PROTEIN.

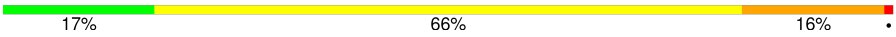
| Mol | Chain | Residues | Atoms | | | | | | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | S | |
| 1 | A | 76 | 1158 | 364 | 566 | 91 | 135 | 2 | 0 |

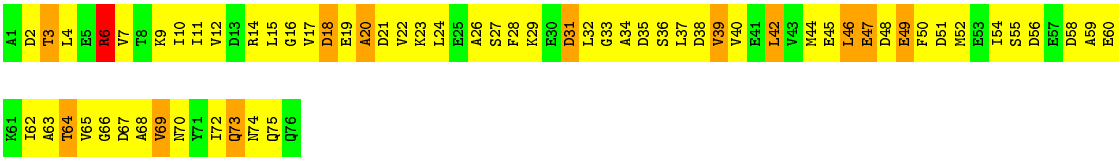
# 4 Residue-property plots (i)

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: ACYL CARRIER PROTEIN

Chain A:



# 5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *Distance geometry, Simulated annealing, molecular dynamics.*

Of the ? calculated structures, 1 were deposited, based on the following criterion: *?*.

The following table shows the software used for structure solution, optimisation and refinement.

| Software name | Classification | Version |
|---|---|---|
| X-PLOR | structure solution | 98.0 |
| X-PLOR | refinement | 98.0 |

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

| Chemical shift file(s) | BMRB entry 4989 |
|---|---|
| Number of chemical shift lists | 1 |
| Total number of shifts | 904 |
| Number of shifts mapped to atoms | 861 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 43 |
| Number of shifts with mapping warnings | 0 |
| Assignment completeness (well-defined parts) | 84% |

No validations of the models with respect to experimental NMR restraints is performed at this time.

# 6 Model quality (i)

## 6.1 Standard geometry (i)

There are no covalent bond-length or bond-angle outliers.

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

| Mol | Chain | Chirality | Planarity |
|---|---|---|---|
| 1 | A | 0 | 2 |
| All | All | 0 | 2 |

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

All planar outliers are listed below.

| Mol | Chain | Res | Type | Group |
|---|---|---|---|---|
| 1 | A | 14 | ARG | Sidechain |
| 1 | A | 6 | ARG | Sidechain |

## 6.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes |
|---|---|---|---|---|---|
| 1 | A | 592 | 566 | 568 | 63 |
| All | All | 592 | 566 | 568 | 63 |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 54.

5 of 63 clashes are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) |
|---|---|---|---|
| 1:A:40:VAL:HG13 | 1:A:59:ALA:HB3 | 1.03 | 1.08 |
| 1:A:40:VAL:HG13 | 1:A:59:ALA:CB | 0.87 | 1.99 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) |
|---|---|---|---|
| 1:A:40:VAL:CG1 | 1:A:59:ALA:HB3 | 0.81 | 2.01 |
| 1:A:11:ILE:HG21 | 1:A:22:VAL:HG23 | 0.76 | 1.57 |
| 1:A:24:LEU:HD23 | 1:A:66:GLY:CA | 0.75 | 2.11 |

## 6.3 Torsion angles (i)

### 6.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles |
|---|---|---|---|---|---|---|
| 1 | A | 74/76 (97%) | 59 (80%) | 10 (14%) | 5 (7%) | 3 18 |
| All | All | 74/76 (97%) | 59 (80%) | 10 (14%) | 5 (7%) | 3 18 |

All 5 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 29 | LYS |
| 1 | A | 33 | GLY |
| 1 | A | 17 | VAL |
| 1 | A | 39 | VAL |
| 1 | A | 20 | ALA |

### 6.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| 1 | A | 66/66 (100%) | 38 (58%) | 28 (42%) | 0 3 |
| All | All | 66/66 (100%) | 38 (58%) | 28 (42%) | 0 3 |

5 of 28 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 70 | ASN |
| 1 | A | 38 | ASP |
| 1 | A | 3 | THR |
| 1 | A | 46 | LEU |
| 1 | A | 27 | SER |

#### 6.3.3 RNA (i)

There are no RNA molecules in this entry.

### 6.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.5 Carbohydrates (i)

There are no carbohydrates in this entry.

### 6.6 Ligand geometry (i)

There are no ligands in this entry.

### 6.7 Other polymers (i)

There are no such molecules in this entry.

### 6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

# 7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 84% for the well-defined parts and 84% for the entire structure.

## 7.1 Chemical shift list 1

File name: BMRB entry 4989

Chemical shift list name: *assigned_chem_shift_list_1*

### 7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

| | |
|---|---|
| Total number of shifts | 904 |
| Number of shifts mapped to atoms | 861 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 43 |
| Number of shifts with mapping warnings | 0 |
| Number of shift outliers (ShiftChecker) | 1 |

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. First 5 (of 43) occurences are reported below.

| Chain | Res | Type | Atom | Shift Data | | |
|---|---|---|---|---|---|---|
| | | | | Value | Uncertainty | Ambiguity |
| A | 3 | LEU | CD1 | 25.05 | -1.0 | 1 |
| A | 2 | PRO | HA | 4.53 | -1.0 | 1 |
| A | 2 | PRO | HB3 | 2.3178 | -1.0 | 1 |
| A | 1 | GLY | HA2 | 4.04 | -1.0 | 1 |
| A | 3 | LEU | HD21 | 0.9565 | -1.0 | 1 |

### 7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

| Nucleus | # values | Correction ± precision, *ppm* | Suggested action |
|---|---|---|---|
| $^{13}C_{\alpha}$ | 81 | -0.64 ± 0.19 | Should be applied |
| $^{13}C_{\beta}$ | 76 | 0.27 ± 0.09 | None needed (< 0.5 ppm) |
| $^{13}C'$ | 0 | — | — |

*Continued on next page...*

*Continued from previous page...*

| Nucleus | # values | Correction ± precision, *ppm* | Suggested action |
|---|---|---|---|
| $^{15}$N | 79 | 0.47 ± 0.33 | None needed (< 0.5 ppm) |

### 7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 84%, i.e. 744 atoms were assigned a chemical shift out of a possible 885. 16 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Backbone | 304/380 (80%) | 152/152 (100%) | 76/152 (50%) | 76/76 (100%) |
| Sidechain | 422/479 (88%) | 256/270 (95%) | 161/194 (83%) | 5/15 (33%) |
| Aromatic | 18/26 (69%) | 12/14 (86%) | 6/12 (50%) | 0/0 (—%) |
| Overall | 744/885 (84%) | 420/436 (96%) | 243/358 (68%) | 81/91 (89%) |

### 7.1.4 Statistically unusual chemical shifts (i)

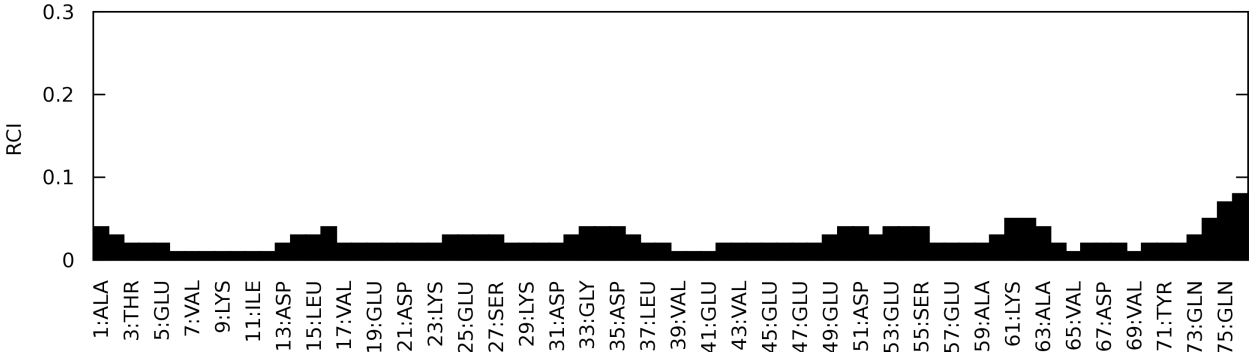
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

| Mol | Chain | Res | Type | Atom | Shift, *ppm* | Expected range, *ppm* | Z-score |
|---|---|---|---|---|---|---|---|
| 1 | A | 49 | GLU | HG3 | 0.99 | 3.31 – 1.21 | -6.0 |

### 7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:


RCI
0.3
0.2
0.1
0
1:ALA
3:THR
5:GLU
7:VAL
9:LYS
11:ILE
13:ASP
15:LEU
17:VAL
19:GLU
21:ASP
23:LYS
25:GLU
27:SER
29:LYS
31:ASP
33:GLY
35:ASP
37:LEU
39:VAL
41:GLU
43:VAL
45:GLU
47:GLU
49:GLU
51:ASP
53:GLU
55:SER
57:GLU
59:ALA
61:LYS
63:ALA
65:VAL
67:ASP
69:VAL
71:TYR
73:GLN
75:GLN